

PARTIAL IDENTIFICATION IN TWO-SIDED MATCHING MODELS

FEDERICO ECHENIQUE, SANGMOK LEE, AND MATTHEW SHUM

In this paper we propose a methodology for estimating preference parameters in matching models. Our estimator applies to repeated observations of matchings among a fixed group of individuals, which is a similar data structure as in Fox (2010). Our estimator is based on stability conditions in the matching models; we consider both transferable (TU) and non-transferable utility (NTU) models. In both cases, the stability conditions yield moment inequalities which can be taken to the data. The preference parameters are partially identified. We consider simple illustrative examples, and also an empirical application to aggregate marriage markets.

1. SETUP

Consider a setup where we observe repeated individual-level matchings among a group of N men and women. Index men (women) by $m = 1, \dots, N$ ($w = 1, \dots, N$). The set of men (women) is denoted M (W). Each man $m \in M$ has a set of strict preferences $>_m$ over $W \cup \{m\}$; similarly, each woman $w \in W$ has a set of strict preferences $>_w$ over men $M \cup \{w\}$. We assume that matchings are one-to-one.

Moment conditions are defined for each potential pair of couples through the stability conditions. First, define the utility indicators, for $i, k \in M$ and $j, l \in W$

$$d_{ijl} := \mathbb{1}(j >_i l) \quad \text{and} \quad d_{jik} := \mathbb{1}(i >_j k).$$

For the **nontransferable utility (NTU) model**, the stability condition implies: for $i, k \in M$ and $j, l \in W$, with $i \neq k$ and $j \neq l$,

$$(1) \quad (i, j), (k, l) \text{ matched} \Rightarrow \begin{cases} d_{ilj}d_{lik} = 0, \text{ and} \\ d_{jki}d_{kjl} = 0 \end{cases}$$

This stability condition implies the moment inequality:

$$Pr((i, j), (k, l) \text{ matched}) \leq Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0).$$

Date: June 20, 2013

Echenique and Shum are affiliated with the Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125; Lee is with Department of Economics, University of Pennsylvania, Philadelphia, PA 19104. Emails: fede@hss.caltech.edu, mshum@caltech.edu, sangmok@sas.upenn.edu.

For the **transferable utility (TU) model**, let $\mathcal{A} = (\alpha_{i,j})$ be a $|M| \times |W|$ matrix of non-negative real numbers. \mathcal{A} is called a *surplus matrix*, in which $\alpha_{i,j}$ is the surplus jointly generated by man- i and woman- j .

A matching is called *optimal* if it achieves the maximum total surplus. It is well-known that optimality corresponds to the appropriate notion of stability for the TU model (Shapley and Shubik, 1971). The formal notion of stability requires a discussion of agents' payoffs; for reasons of space, we omit the definition of stability and focus instead on optimal matchings.

A *necessary condition* of the optimal matching is the *pairwise-stability* condition:

$$(i, j), (k, l) \text{ matched} \Rightarrow \alpha_{ij} + \alpha_{kl} \geq \alpha_{il} + \alpha_{kj}.$$

This leads to the moment inequality

$$Pr((i, j), (k, l) \text{ matched}) \leq Pr(\alpha_{ij} + \alpha_{kl} \geq \alpha_{il} + \alpha_{kj}; \beta).$$

For both NTU and TU models, the LHS of the moment inequalities can be obtained directly from the data, as sample frequencies when the number of repeated matchings grows large. The RHS of the moment inequalities will depend on the utility parameters, once we specify the utility functions. (A simple example will be presented in the next section.) The number of moment conditions then is the number of potential pairs of couples that can be observed; out of N men and women, there are N^2 potential couples that can be formed; hence, there are $N^2 * (N - 1)^2$ *pairs* of potential couples consisting of two distinct men and two distinct women.

1.1. Comparison with other estimation approaches. Note that generally, parameters in both NTU and TU setting will be partially identified. For the NTU setting, this is due to the multiplicity of stable matchings, and echoes the partial identification results for game models with multiple equilibria (cf. Ciliberto and Tamer (2009), Beresteanu et al. (2011)). For the TU setting, even though the optimal matching is generally unique, we are using only *necessary conditions* for identification, and hence partial identification results. This contrasts with Fox (2010), who considers maximum score estimation of the TU model using the pairwise stability conditions, and obtains point identification of utility parameters.

2. EXAMPLE: NTU MODEL

Here we present a simple 2×2 example with two men (i, k) and two women (j, l) . Utilities are:

$$\begin{aligned} U_{m,w} &= \beta_M |Age_m - Age_w| + \varepsilon_{m,w}; \\ U_{w,m} &= \beta_W |Age_m - Age_w| + \varepsilon_{w,m}. \end{aligned}$$

Utilities depend just on the age differences between the matched persons. The unobserved portion of utility, ε , is assumed to be distributed i.i.d. $N(0, \frac{1}{2})$ across all m, w , and identically for men and women.

With just two men and two women, there are only two pairs of distinct potential couples: $\{(i, j), (k, l)\}$ and $\{(i, l), (k, j)\}$. Hence there are two moment inequalities. Assume that, from the data, we have the following match frequencies:

$$\begin{aligned} Pr(\{(i, j), (k, l)\}) &= 0.3 \\ Pr(\{(i, l), (k, j)\}) &= 1 - Pr(\{(i, j), (k, l)\}) = 0.7. \end{aligned}$$

We consider the NTU model here. Hence the first moment inequality says

$$Pr(\{(i, j), (k, l)\}) \leq Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0).$$

Given the utility specification above, this becomes (letting $\Delta_{ij} := |Age_i - Age_j|$):

$$\begin{aligned} Pr(\{(i, j), (k, l)\}) &\leq [1 - \Phi(\beta_M(\Delta_{il} - \Delta_{ij}))\Phi(\beta_W(\Delta_{il} - \Delta_{kl}))] \\ &\quad \cdot [1 - \Phi(\beta_W(\Delta_{jk} - \Delta_{ji}))\Phi(\beta_M(\Delta_{kj} - \Delta_{kl}))]. \end{aligned}$$

Analogously, for the second moment inequality, we have:

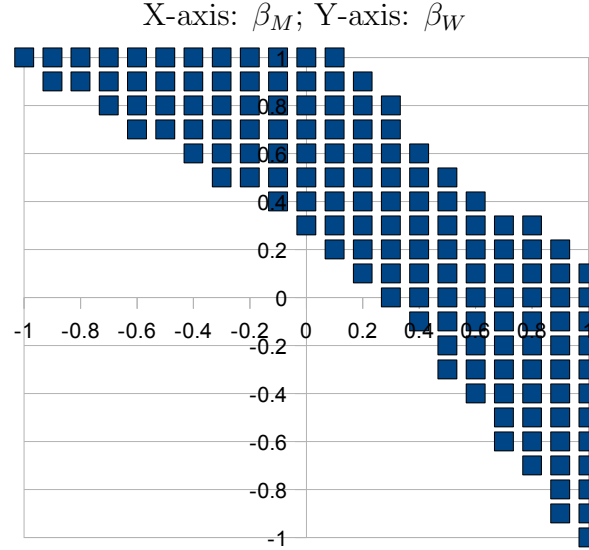
$$\begin{aligned} 1 - Pr(\{(i, j), (k, l)\}) &\leq [1 - \Phi(\beta_M(\Delta_{kl} - \Delta_{kj}))\Phi(\beta_W(\Delta_{kl} - \Delta_{il}))] \\ &\quad \cdot [1 - \Phi(\beta_W(\Delta_{ij} - \Delta_{kj}))\Phi(\beta_M(\Delta_{ij} - \Delta_{il}))]. \end{aligned}$$

For our example, we have that $\Delta_{kj} = \Delta_{il} = 0.5$, while $\Delta_{ij} = \Delta_{kl} = 0$.

The identified set for the NTU version of this simple example is shown in Figure 1. To a certain degree, the admissible preferences of men and women have an “antipodal” feature. When $\beta_M \ll 0$, implying that men dislike a large age gap, then $\beta_W \gg 0$, implying that women prefer a larger age gap. When men prefer a larger age gap ($\beta_M > 0$), however, then women may be either indifferent or dislike a large age gap ($\beta_W \leq 0$).

Such antipodal preferences is consistent with a general logic of stability in NTU settings. In such a setting, an observed matching may not be indicative that each person is matched to a “most preferred” partner; rather, stability of the observed matching only implies that (say) each man is not able to find a more preferable partner

FIGURE 1. Identified set for simple 2x2 example: NTU model



who would also prefer to be matched with him rather than her current mate. It is this “no blocking pairs” requirement which restricts the utility parameter values. In the example, we assumed that “unequal-aged matchings” (i.e. $\{(k, j), (i, l)\}$) occurred with probability 0.7, which is higher than the “equal-age matchings” (i.e. $\{(i, j), (k, l)\}$). The “no blocking pairs” conditions then implies that if men value an equal-aged partner (i.e. $\beta_M < 0$), then women must value an unequal-aged partner (i.e. $\beta_W > 0$), and vice-versa. We see these implications in the identified set of admissible preference parameters.

3. AGGREGATE MATCHING MODEL

One problem with the present estimator is the need to observe repeated matchings from comparable populations. Also, the population should be relatively small, in order for the number of moment conditions to stay modest and manageable. Both of these requirements are difficult to fulfill in practice. Therefore, in this section we consider the robustness of our estimator when applied to *aggregate* data: that is, when the data available are tables of the match frequencies for different aggregate *types* of agents. We spell out the theory of such aggregate matchings in another paper (Echenique et al. (2013)). Here we introduce and define basic concepts, which will be used in the empirical application below.

An *aggregate matching market* is described by a triple $\langle M, W, > \rangle$, where:

- (1) M and W are disjoint, finite sets.

- (2) $> := ((>_m)_{m \in M}, (>_w)_{w \in W})$ is a profile of strict preferences: for each m and w , $>_m$ is a linear order over $W \cup \{m\}$ and $>_w$ is a linear order over $M \cup \{w\}$.

We call agents on one side men, and on the other side women, as is traditional in the matching literature. The elements of M are *types of men*, and the elements of W are *types of women*. Many applications are, of course, to environments different from the marriage matching market.

Note that preferences $>$ above effectively rules out preference heterogeneity among agents of the same type. While this is restrictive relative to other aggregate matching models in the literature, such as Choo and Siow (2006), Galichon and Salanie (2009), both of these papers consider the TU model. For the NTU model (which is the focus of this section), stability conditions for a model with agent-specific preference heterogeneity has *no empirical implications* at the aggregate level (see Appendix A for further discussions). For this reason, we assume that all agents of the same type have identical preferences.

Consider an aggregate matching market $\langle M, W, > \rangle$, with $M = \{m_1, \dots, m_K\}$ and $W = \{w_1, \dots, w_L\}$. An *aggregate matching* is a $K \times L$ matrix $X = (X_{ij})$ with non-negative integer entries. The interpretation of X is that X_{ij} is the number of type- i men and type- j women matched to each other. An aggregate matching X is *canonical* if $X_{ij} \in \{0, 1\}$. For any aggregate matching X , we can construct a canonical aggregate matching X^c by setting $X_{ij}^c = 0$ when $X_{ij} = 0$ and $X_{ij}^c = 1$ when $X_{ij} > 0$.

We consider, in turn, the nontransferable utility model and its empirical implementation, followed by the transferable utility model.

3.1. Non-transferable utility model. An aggregate matching X is *stable* if it is individually rational and there are no blocking pairs for X . Obviously, an aggregate matching X is stable if and only if the corresponding canonical matching X^c is stable. Therefore, our empirical results below pertain to canonical aggregate matchings.

Given a canonical matching X , we define an *anti-edge* as a pair of couples $\{(i, j), (k, l)\}$ with $i \neq k \in M$ and $j \neq l \in W$ such that $X_{ij} = X_{kl} = 1$. Then, stability of the canonical aggregate matching X is equivalent to:

$$(2) \quad (i, j), (k, l) \text{ anti-edge} \Rightarrow \begin{cases} \mathbb{1}(w_l >_{m_i} w_j) \cdot \mathbb{1}(m_i >_{w_l} m_k) = 0, \text{ and} \\ \mathbb{1}(w_j >_{m_k} w_l) \cdot \mathbb{1}(m_k >_{w_j} m_i) = 0. \end{cases}$$

In our empirical work with the NTU model, Eq. (2) of the stability conditions forms the basis for the moment inequalities. The anti-edge condition (2) implies that

$$(3) \quad Pr((i, j), (k, l) \text{ anti-edge}) \leq Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0).$$

Given parameter values β , and our assumptions regarding the distribution of the ε 's, these probabilities can be calculated. Hence, the moment inequality corresponding to Eq. (3) is:

$$(4) \quad \mathbb{E} \underbrace{[\mathbb{1}((i, j), (k, l) \text{ anti-edge}) - Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0; \beta)]}_{g_{ijkl}(X; \beta)} \leq 0.$$

The identified set is defined as

$$\mathbb{B}_0 = \{\beta : \mathbb{E}g_{ijkl}(X; \beta) \leq 0, \forall i, j, k, l\}.$$

These moment inequalities are quite distinct from the estimating equations considered in the existing empirical matching literature. For instance, Choo and Siow (2006), Dagsvik (2000), and Fox (2010) use equations similar to those in the multinomial choice literature, that each observed pair (i, j) represents, for both i and j , an “optimal choice” from some “choice set”. The restrictions in (2) cannot be expressed in such a way.

Assume that we observe multiple aggregate matchings. Let T be the number of such observations, and X_t denote the t -th aggregate matching that we observe. Then the sample analog of the expectation in (4) is

$$(5) \quad \begin{aligned} & \frac{1}{T} \sum_t \mathbb{1}((ij), (kl) \text{ is anti-edge in } X_t) - Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0; \beta) \\ &= \frac{1}{T} \sum_t g_{ijkl}(X_t; \beta). \end{aligned}$$

If the number of types of men and woman were equal ($K = L$), then there would be $\frac{K^2 \times (K-1)^2}{2}$ such inequalities, corresponding to each couple of pairs. Note that the expectation \mathbb{E} above is over both the utility shocks ε 's, as well as over the “equilibrium selection” process (which we are agnostic about).

There is by now a large methodological literature on estimating confidence sets for parameters in partially identified moment inequality models that cover the identified set \mathbb{B}_0 with some prescribed probability. (An incomplete list includes Chernozhukov et al. (2007), Andrews et al. (2004), Romano and Shaikh (2010), Pakes et al. (2007), Beresteanu and Molinari (2008).) While there are a variety of objective functions one could use, we use here the simple sum of squares objective:

$$\mathbb{B}_n = \operatorname{argmin}_{\beta} Q_n(\beta) = \sum_{i,j,k,l} \left[\frac{1}{T} \sum_{t=1}^T g_{ijkl}(X_t; \beta) \right]_+^2$$

where $[x]_+ := \max\{x, 0\}$.

3.2. Data and empirical implementation. In the empirical implementation, we use data on new marriages, as recorded by the US Bureau of Vital Statistics.¹ We consider new marriages in the year 1988, and treat data from each state as a separate, independent matching. We aggregate the matchings into age categories, and create canonical matchings. For this application, we only include the age variable in our definition of agent types, because it is the only variable which we observe for all the matchings.² Table 1 has examples of aggregate matchings, and the corresponding canonical matchings, for several states. In these matching matrices, rows denote age categories for the husbands, and the columns denote the age categories for the wives.

TABLE 1. Aggregate Matchings and the corresponding Canonical Matchings.

	Age: $\sigma \downarrow, \varphi \rightarrow$	Aggregate Matchings							Canonical Matchings						
		12-20	21-25	26-30	31-35	36-40	41-50	51-94	12-20	21-25	26-30	31-35	36-40	41-50	51-94
MI	12-20	231	47	8	0	0	1	0	1	1	1	0	0	1	0
	21-25	329	798	156	32	11	7	0	1	1	1	1	1	1	0
	26-30	71	477	443	136	27	8	0	1	1	1	1	1	1	0
	31-35	11	148	249	196	83	21	0	1	1	1	1	1	1	0
	36-40	2	41	105	144	114	51	1	1	1	1	1	1	1	1
	41-50	0	15	42	118	121	162	25	0	1	1	1	1	1	1
	51-94	0	2	11	11	35	137	158	0	1	1	1	1	1	1
NV	12-20	8	1	0	0	0	0	0	1	1	0	0	0	0	0
	21-25	17	31	4	0	0	0	0	1	1	1	0	0	0	0
	26-30	2	21	22	7	1	0	0	1	1	1	1	1	0	0
	31-35	0	4	10	5	3	0	0	0	1	1	1	1	0	0
	36-40	0	3	8	2	2	2	0	0	1	1	1	1	1	0
	41-50	0	1	1	2	6	3	3	0	1	1	1	1	1	1
	51-94	0	0	0	0	0	5	3	0	0	0	0	0	1	1
PA	12-20	307	83	12	6	0	0	0	1	1	1	1	0	0	0
	21-25	453	1165	214	64	10	6	1	1	1	1	1	1	1	1
	26-30	113	698	703	190	51	17	0	1	1	1	1	1	1	0
	31-35	17	184	393	277	78	26	2	1	1	1	1	1	1	1
	36-40	9	73	152	191	148	84	5	1	1	1	1	1	1	1
	41-50	3	27	83	146	187	273	28	1	1	1	1	1	1	1
	51-94	1	7	12	38	48	182	268	1	1	1	1	1	1	1

These aggregate canonical matchings have many 1's, and hence many anti-edges. Moreover, the matchings in Table 1 contain more non-zero entries below the diagonal, which means that in a preponderance of marriages, the husband is older than the wife.

In our empirical exercise, the specification of utility is very simple, and it only involves the ages of the two partners to a match. Suppose that man m of age Age_m is matched

¹<http://www.nber.org/data/marrdivo.html>

²Because stability is defined at the level of the matching, we did not want to exclude any marriage from the data due to missing variables.

to woman w of age Age_w . The following utility functions capture preferences over age differences, and partner's age.

$$\begin{aligned} U_{m,w} &= \beta_1 |Age_m - Age_w|^- + \beta_2 |Age_m - Age_w|^+ + \varepsilon_{m,w} \\ U_{w,m} &= \beta_3 |Age_m - Age_w|^- + \beta_4 |Age_m - Age_w|^+ + \varepsilon_{w,m}, \end{aligned}$$

where $\varepsilon_{m,w}$ and $\varepsilon_{w,m}$ are assumed to follow a standard normal distributions.

In this specification, we assume that utility is a piecewise-linear function of age, with the “kink” occurring when the age-gap between husband and wife is zero. To interpret the preference parameters, note that β_1 (β_3) is the coefficient in the husband's (wife's) utility, attached to the age gap when the wife is older than the husband. Thus, a finding that $\beta_1(\beta_3) > 0$ means that, when the wife older, men (women) prefer a larger age gap: that is, men prefer older women, and women prefer younger men. Similarly, a finding that $\beta_2(\beta_4) > 0$, implies that then when the husband is older than the wife, men (women) prefer a larger age gap: here, because the husband is older, a larger age gap means that men prefer younger women, and women prefer older men.

3.3. Relaxing the stability constraints. Stability (rationalizability) places very strong demands on the data that can be observed, since we often observe many 1's, and hence many anti-edges, in aggregate canonical matchings (See Eq. (2) and (3)). Accordingly, we propose a relaxation of the stability constraint that is particularly useful in applied empirical work.

Namely, we assume that potential blocking pairs may not necessarily form. If preferences are such that the pair (m, w) would block X , the block actually occurs only with probability less than 1. The reason for not blocking could be simply the failure of m and w to meet or communicate (as in the literature on search and matching).

Specifically, we allow for the possibility that an observed edge between pairs (i, j) and (k, l) may imply nothing about the preferences of the affected types i, j, k, l , simply because the couples (i, j) and (k, l) fail to meet. In particular, define

$$\delta_{ijkl} = Pr(\text{types } (i, j), (k, l) \text{ communicate}).$$

We then modify the stability inequalities (2) as:

$$(6) \quad \begin{pmatrix} (i, j), (k, l) \text{ is anti-edge} \\ (i, j), (k, l) \text{ meet} \end{pmatrix} \Rightarrow \begin{cases} d_{ilj}d_{lik} = 0 \\ d_{jki}d_{kjl} = 0. \end{cases}$$

This leads to the modified moment inequality:

$$(7) \quad Pr((i, j), (k, l) \text{ anti-edge}) \leq \frac{Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0; \beta)}{\delta_{ijkl}}.$$

As $\delta_{ijkl} \rightarrow 1$, the identified set \mathbb{B}_0 shrinks to the empty set. Thus the observed aggregate matchings cannot be rationalized without a positive probability that potential blocking pairs do not form. On the other hand, as $\delta_{ijkl} \rightarrow 0$, the identified set converges to the whole parameter space: the right-hand side of the moment inequality becomes larger than 1.

Here, the events $((i, j), (k, l) \text{ is anti-edge})$ and $((i, j), (k, l) \text{ meet})$ are independent. The first event depends on preferences and process that produces a stable matching in the first place. On the other hand, we allow δ_{ijkl} to depend on the relative number of matched (i, j) and (k, l) couples. So we are making the assumption that the probability of communication is independent of preferences and the matching.³ Specifically, letting γ denote a scaling parameter, we set

$$\delta_{ijkl} = \min\left\{ 2 \cdot \gamma \cdot \frac{X_{ij}}{N} \cdot \frac{X_{kl}}{N}, 1 \right\}$$

where N is the number of observed men (women).

To interpret this, consider a given pair of couples $(i, j), (k, l)$. If this couple constitutes an anti-edge, and the stability conditions fails, then two potential blocking pairs can be formed: (i, l) and (k, j) . The specification for δ_{ijkl} represents one story for when a blocking pair which is present in the agents' preferences, actually blocks. With X_{ij}/N (resp. X_{kl}/N) being the relative populations of (i, j) (resp. (k, j)) couples, then δ_{ijkl} is set proportional to the frequency of potential blocking pairs $(j, l), (k, j)$ in the market; it is scaled by γ (and capped from above by 1). We scale by γ to allow the probability that a blocking pair forms to be smaller or larger than this frequency. A larger γ implies that blocking pairs form more frequently, so that there is less slackness in the stability restrictions.

More broadly, the δ s weight the anti-edges in the sample moment inequalities. Intuitively, an anti-edge $\{(i, j), (l, k)\}$ should receive a higher weight when it involves many potential blocking pairs than when it only involves a few. Our specification achieves this idea, as it makes the probability of forming a blocking pair dependent on the number of agents involved.

3.4. Identified sets. Table 2 summarizes the identified set for several levels of γ , and presents the highest and lowest values that each parameter attains in the identified set. The unrestricted interval in which we searched for each parameter was $[-2, 2]$. So we see that, for a value of $\gamma = 25$, the identified set contains the full parameter space, implying that the data impose no restrictions on parameters. At the other extreme,

³We could relax this assumption by making δ dependent on the same covariates that enter into the agents preferences.

TABLE 2. Unconditional Bounds of β .

	β_1		β_2		β_3		β_4	
γ	min	max	min	max	min	max	min	max
25	-2.00	2.00	-2.00	2.00	-2.00	2.00	-2.00	2.00
28	-2.00	1.60	-2.00	2.00	-2.00	1.60	-2.00	2.00
29	-2.00	0.40	-2.00	1.80	-2.00	0.40	-2.00	1.80
30	-2.00	-0.80	-2.00	0.60	-2.00	-0.85	-2.00	0.60

when $\gamma \geq 31$, the identified set becomes empty, implying that the observed matchings can no longer be rationalized.

For $\gamma = 30$, we see that β_1 and β_3 take negative values, while the values of β_2 and β_4 tend to take negative values but also contain small positive values. This suggests that husbands' utilities are decreasing in the wife's age when the wife is older, but when the wife is younger, his utility is less responsive to the wife's age. A similar picture emerges for wives' utilities, which are increasing in the husband's age when the husband is younger, but when the husband is older, the wife's utility is less responsive to her husband's age. All in all, our findings here support the conclusion that husbands' and wives' utilities are more responsive to the partner's age when the wife is older than the husband.

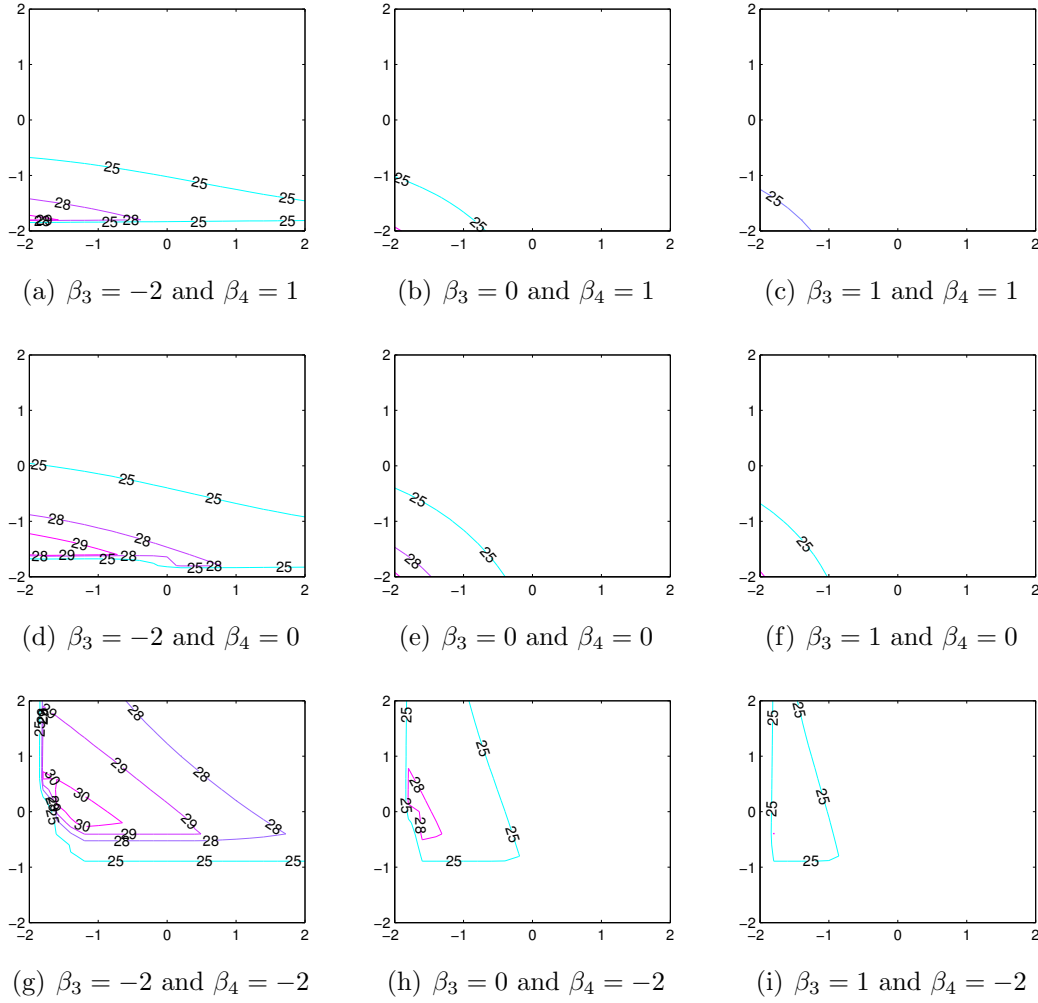
A richer picture emerges when we consider the joint values of parameters in the identified set. Figure 2 illustrates the contour sets (at different values of γ) for the husband's preference parameters (β_1, β_2) , holding the wife's preference parameters (β_3, β_4) fixed.

To simplify the interpretation of these findings in light of the stability restrictions, we recall two features of our aggregate matchings (as seen in Table 1): first, there are more anti-edges below the diagonal, where $Age_m > Age_w$. Second, there are more “downward-sloping” anti-edges than “upward-sloping” ones. That is, there are more anti-edges $\{(i, j), (k, l)\}$ with $k > i, l > j$ than with $i > k, l > j$, as illustrated here.



Because of these features, we initially focus on the parameters (β_2, β_4) , which describe preferences when the husband is older than the wife.

The graphs in the bottom row of Figure 2 correspond to $\beta_4 = -2$, corresponding to the case that the wife prefers a younger husband: with a downward-sloping anti-edge, this implies that it is likely that $d_{jik} = 1$ and $d_{lki} = 0$. In turn, using the stability

FIGURE 2. Identified sets of (β_1, β_2) given (β_3, β_4) and γ .

restrictions (2), this implies that $d_{ilj} = 0$ (that husbands prefer younger wives), but places no restrictions on the sign of d_{kjl} . For this reason, we find that in these graphs, β_2 tends to take positive values at the highest contour levels so that, when husbands are older than their wives, they prefer the age gap to be as large as possible.

By a similar reasoning, β_2 takes negative values when $\beta_4 = 1$. When wives prefer older husbands (which is the case when $\beta_4 = 1$), then with a downward-sloping anti-edge, this implies that $d_{jik} = 0$ and $d_{lki} = 1$. Consequently, stability considerations would restrict the husband's preferences so that $d_{kjl} = 0$ (and husbands prefer older wives), leading to $\beta_2 < 0$.

On the other hand, because there are more downward-sloping anti-edges, when the wife is older than the husband, restriction (2) implies that one of two cases – either the husband prefers a younger wife, or the wife prefers an older husband – must be

true. In Figure 2, as β_3 increases from -2 to 1 (from the left to the right column), the wife's utilities becomes more favorable towards a younger husband. As a result, more restrictions are imposed to the husbands' utilities, which yields a tighter negative range for β_1 in the identified sets.

Overall, we see that $\beta_1 < 0$ and $\beta_3 < 0$, implying that as long as the wife is older than the husband, both prefer a smaller age gap. On the other hand, β_2 and β_4 are negatively correlated: as β_4 increases, β_2 decreases. This suggests that, when the husband is older than the wife, one side prefers a smaller gap but the other side is less responsive on the age gap.

3.5. Confidence sets. Figure 3 summarizes the 95% confidence sets with $\gamma = 28$ (shaded lightly) and 30 (shaded darkly). In computing these confidence sets, we use the subsampling algorithm proposed by Chernozhukov et al. (2007). Comparing the confidence sets in Figure 3 to their counterpart identified sets in Figure 2, the confidence sets are slightly larger than the identified sets. This is not surprising, given the modest number of matchings (fifty-one: one for each state) which we used in the empirical exercise.

Nevertheless, the main findings from Figure 1 are still apparent; $\beta_1 < 0$ across a range of values for (β_3, β_4) , and $\beta_2 < 0$ (resp. > 0) when $\beta_4 > 0$ (resp. < 0). These somewhat “antipodal” preferences between a husband and wife are a distinctive consequence of the stability conditions of an NTU matching model.

3.6. Transferable utility model. For the TU model, we define the surplus obtained by matching of type- i man with type- j woman as:

$$\begin{aligned}\alpha_{ij} &= U_{ij} + U_{ji} \\ &= (\beta_1 + \beta_3)|Age_m - Age_w|^- + (\beta_2 + \beta_4)|Age_m - Age_w|^+ + \varepsilon_{ij} + \varepsilon_{ji}.\end{aligned}$$

We work from the *pairwise stability* condition: for every anti-edge $\{(i, j), (k, l)\}$, we have

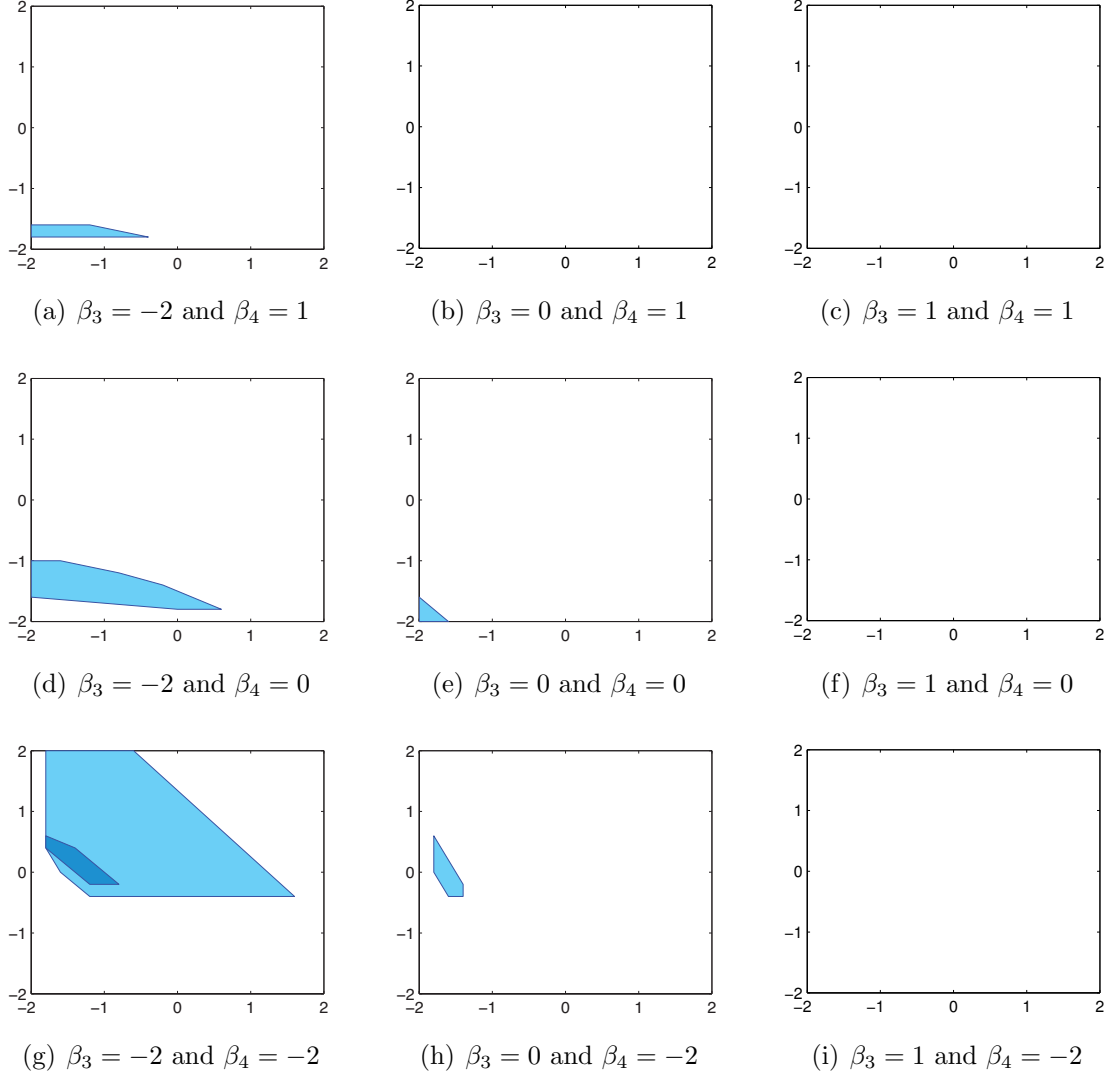
$$(i, j), (k, l) \text{ anti-edge} \Rightarrow \alpha_{ij} + \alpha_{kl} \geq \alpha_{il} + \alpha_{kj}.$$

This leads to the moment inequality

$$Pr((i, j), (k, l) \text{ anti-edge}) \leq Pr(\alpha_{ij} + \alpha_{kl} \geq \alpha_{il} + \alpha_{kj}; \beta).$$

This condition derived via optimality. Given an aggregate matching X , suppose $\{(i, j), (k, l)\}$ is an anti-edge (i.e., $X_{ij} > 0$ and $X_{kl} > 0$). Consider an alternative

FIGURE 3. 95% confidence sets of (β_1, β_2) given (β_3, β_4) and $\gamma = 32$ (shaded lightly) and 35 (shaded darkly).



aggregate matching X' where a pair of $\{(i, j), (k, l)\}$ couples are swapped:

$$\begin{aligned} X'_{ij} &= X_{ij} - 1, & X'_{kl} &= X_{kl} - 1, \\ X'_{il} &= X_{il} + 1, & X'_{kj} &= X_{kj} + 1. \end{aligned}$$

By optimality of X , this swapping must lower surplus:

$$\begin{aligned} \alpha_{ij}X_{ij} + \alpha_{il}X_{il} + \alpha_{kj}X_{kj} + \alpha_{kl}X_{kl} &\geq \alpha_{ij}X'_{ij} + \alpha_{il}X'_{il} + \alpha_{kj}X'_{kj} + \alpha_{kl}X'_{kl} \\ &= \alpha_{ij}(X_{ij} - 1) + \alpha_{il}(X_{il} + 1) \\ &\quad + \alpha_{kj}(X_{kj} + 1) + \alpha_{kl}(X_{kl} - 1) \end{aligned}$$

$$\implies \alpha_{ij} + \alpha_{kl} \geq \alpha_{il} + \alpha_{kj}.$$

For the same reason as in the NTU model, we relax the stability constraints by introducing communication probabilities:

$$Pr((i, j), (k, l) \text{ anti-edge}) \leq \frac{Pr(\alpha_{ij} + \alpha_{kl} \geq \alpha_{il} + \alpha_{kj}; \beta)}{\delta_{ijkl}}.$$

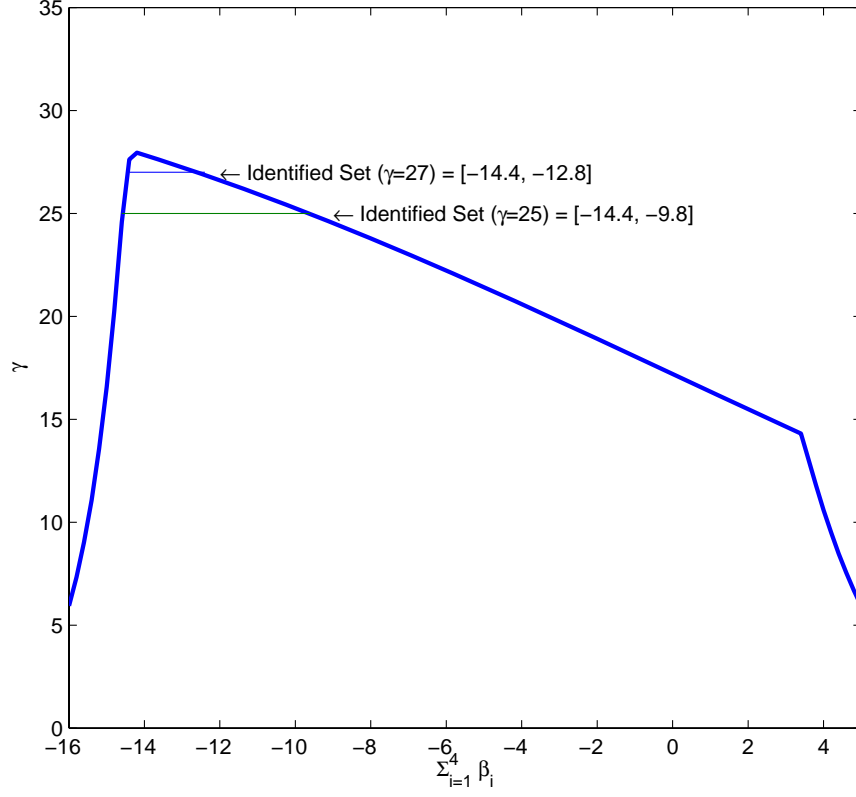
The identified set for the TU model takes the form $K_1 \leq \sum_{i=1}^4 \beta_i \leq K_2$. First, if all (i, j) , (i, l) , (k, j) , and (k, l) are “below the diagonal” (i.e. $Age_m > Age_w$):

$$\begin{aligned} \alpha_{ij} + \alpha_{kl} &= (\beta_2 + \beta_4)|Age_i - Age_j|^+ + (\beta_2 + \beta_4)|Age_k - Age_l|^+ + \Sigma_{ij,kl} \\ &\quad (\text{since } Age_i > Age_j \text{ and } Age_k > Age_l) + \Sigma_{ij,kl} \\ &= (\beta_2 + \beta_4)(Age_i - Age_j) + (\beta_2 + \beta_4)(Age_k - Age_l) + \Sigma_{ij,kl} \\ &= (\beta_2 + \beta_4)(Age_i - Age_l) + (\beta_2 + \beta_4)(Age_k - Age_j) + \Sigma_{ij,kl} \\ &= \alpha_{il} + \alpha_{kj} - \Sigma_{il,kj} + \Sigma_{ij,kl} \\ &\quad (\text{since } Age_i > Age_l \text{ and } Age_k > Age_j) \end{aligned}$$

where we define the shorthand $\Sigma_{ij,kl} = \varepsilon_{ij} + \varepsilon_{ji} + \varepsilon_{kl} + \varepsilon_{lk}$. Since the event $\alpha_{ij} + \alpha_{kl} \geq \alpha_{il} + \alpha_{kj}$ is equivalent to $\Sigma_{ij,kl} \geq \Sigma_{il,kj}$, and involves no model parameters, stability (rationalizability) imposes no restriction on the data in the “below diagonal” case. Similarly, stability imposes no restriction on the observed matchings for (i, j) , (k, l) , (i, l) , and (k, j) which are all above diagonal (i.e., $Age_m < Age_w$).

Therefore, identification is determined by moment conditions corresponding to men i and k , and women j and l , where we have a pair below diagonal, and a pair above diagonal. Suppose, for example, (i, j) , (k, l) , and (k, j) are below diagonal, but (i, l) is above diagonal (i.e., $Age_k > Age_l > Age_i > Age_j$):

$$\begin{aligned} \alpha_{ij} + \alpha_{kl} &= (\beta_2 + \beta_4)|Age_i - Age_j|^+ + (\beta_2 + \beta_4)|Age_k - Age_l|^+ + \Sigma_{ij,kl} \\ &\quad (\text{since } Age_i > Age_j \text{ and } Age_k > Age_l) \\ &= (\beta_2 + \beta_4)(Age_i - Age_j) + (\beta_2 + \beta_4)(Age_k - Age_l) + \Sigma_{ij,kl} \end{aligned}$$

FIGURE 4. Identified sets of $\sum_{i=1}^4 \beta_i$ given γ .

$$\begin{aligned}
& \alpha_{il} + \alpha_{kj} \\
&= (\beta_1 + \beta_3)|Age_i - Age_l|^- + (\beta_2 + \beta_4)|Age_k - Age_j|^+ + \Sigma_{il,kj} \\
&\quad (\text{since } Age_i < Age_l \text{ and } Age_k > Age_j) \\
&= (\beta_1 + \beta_3)(Age_l - Age_i) + (\beta_2 + \beta_4)(Age_k - Age_j) + \Sigma_{il,kj}.
\end{aligned}$$

Therefore,

$$(\alpha_{ij} + \alpha_{kl}) - (\alpha_{il} + \alpha_{kj}) = (\beta_1 + \beta_2 + \beta_3 + \beta_4)(Age_i - Age_j) + \Sigma_{ij,kl} - \Sigma_{il,kj}$$

For all other cases, we have the same result: we can identify β up to $\sum_{i=1}^4 \beta_i$.

The identified set is presented in Figure 4, which is consistent with both antipodal and non-antipodal preferences.

4. CONCLUSIONS

In this paper we propose a methodology for estimating preference parameters in matching models. Our estimator applies to repeated observations of matchings among a fixed group of individuals. For both the transferable utility (TU) and non-transferable utility (NTU) models, we derive moment inequalities based on the restrictions which match stability places on the preferences of the agents.

APPENDIX A. INDIVIDUAL-LEVEL HETEROGENEITY

In our theoretical results, we have assumed that agents' preferences depend only on observables. This allowed us to obtain rather stark implications of stability for aggregate matchings. Maybe the implications are too stark, in the sense that most of the observed matchings in the data would not be rationalizable. If we add unobserved heterogeneity, then the theoretical implications become weaker and “probabilistic”, but the main thrust of these implications are preserved.

So, in a matching model that captures how preferences depend on observables, but has additional noise, our conditions for rationalizability hold in a probabilistic sense. The econometric approach proposed here involves just such a probabilistic version of the model. Here we compare our approach to other papers in the literature.

One possible starting point is to assume that individuals of the same type have the same preferences up to individual-specific i.i.d. shocks, which is the assumption in most of the empirical literature.⁴

The i.i.d. shocks are a very limited form of unobserved heterogeneity: it allows two (say) type i men to differ in the utility they would obtain from a matching with a (say) type j woman. However, each of these men still remains indifferent between all type j women.⁵ Thus two agents of the same type are still perceived as identical by the opposite side of the market.

The shocks ensure that each agent-type has a non-zero probability of being matched with any agent-type on the opposite side of the market; this reconciles the theory with the observed data. In this respect, the role of the preference shocks in these papers plays the same role as the “communication probability” δ_{ijkl} in our empirical analysis. The “communication probability” captures unobserved heterogeneity in the ability of agents to match, perhaps as a result of noisy search frictions. It serves the same purpose as i.i.d. preference shocks. The shocks, on the other hand, lead to trivial inequalities at the aggregate level. We state this result here, and prove it in the Appendix B.

Claim 1. *In the NTU model, preference shocks at the individual-level lead to trivially-satisfied stability restrictions at the aggregate level.*

Because of this result, then, i.i.d. individual-level preference shocks seem inappropriate in the aggregate NTU setting of our empirical work. Furthermore, the communication probability δ_{ijkl} plays a similar role in our empirical work as do preference shocks

⁴See, for instance, Choo and Siow (2006) and Galichon and Salanie (2009) for the TU model. For NTU model, Uetake and Watanabe (2012) takes this utility specification with an assumption that the model generates unique stable matchings.

⁵Galichon and Salanie (2009) also discuss this point (cf. pg. 10).

in others' work: namely, to better reconcile the theory to the data by enlarging the the sets of marriages which one could observe in a stable matching.

The sample moment inequality (Eq. (5)), with the modification in Eq. (6), becomes:

$$\begin{aligned} & \frac{1}{T} \sum_t g_{ijkl}(X_t; \beta) \\ &= \left(\frac{1}{T} \sum_t \mathbb{1}((i, j), (k, l) \text{ anti-edge in } X_t) \right) * \delta_{ijkl} - \Pr(d_{ilj}d_{lik} = 0, d_{jki}d_{kjl} = 0; \beta) \end{aligned}$$

for all combinations of pairs (i, j) and (k, l) .

APPENDIX B. DETAILS ON CLAIM 1

We consider a market where every woman (man) is acceptable to all men (women). The individual-level stability inequalities, for all pairs (i, j) , are:

$$\sum_{k:k>i,j} X_{ik} + \sum_{k:k>j,i} X_{kj} + X_{ij} \geq 1.^6$$

Letting $d_{ikj} = \mathbb{1}\{k >_i j\}$, this can be written as:

$$(8) \quad \sum_k X_{ik} d_{ikj} + \sum_k X_{kj} d_{jki} + X_{ij} \geq 1.$$

Here (i, j, k) all denote individual agents, not types. These inequalities cannot be taken directly to the data, because we do not observe the individual-level matching, but rather an aggregate-level matching.

One starting point is to treat both the X 's and the d 's as random variables, where the randomness derives from both the individual-level preference shocks, as well as from the procedure whereby the observed matching is selected among the set of stable matchings. We partition the men and women into types t_1^M, \dots, t_L^M and t_1^W, \dots, t_L^W . Since individual-level preference shocks are i.i.d., we obtain that

$$(9) \quad \Pr(d_{ijk} = 1) = \Pr(d_{i'j'k'} = 1) : \quad \forall (i, i') \in t_i^M, (j, j') \in t_j^M, (k, k') \in t_k^M.$$

That is, the distribution of d_{ijk} is identical for all individuals of the same type. Hence, below we will use the notation $\Pr(d_{ijk} = 1)$ and $\Pr(t_j^W >_{t_i^M} t_k^W)$ interchangeably.

⁶These individual-level inequalities express the same notion of stability as the aggregate stability conditions (2), but can be written in this more succinct way here due to the summing-up requirements at the individual-level (i.e., that $\sum_j X_{ij} = 1$ for all i). These summing-up conditions do not hold for canonical aggregate matchings.

Given these assumptions, we can derive an aggregate version of Eq. (8). First, we take expectations:

$$\begin{aligned} \sum_k E[X_{ik}d_{ikj}] + \sum_k E[X_{kj}d_{jki}] + E[X_{ij}] &\geq 1 \\ \iff \sum_k \bar{X}_{ikj} \cdot Pr(d_{ikj} = 1) + \sum_k \bar{X}_{kji} \cdot Pr(d_{jki} = 1) + E[X_{ij}] &\geq 1 \end{aligned}$$

with $\bar{X}_{ikj} \equiv E[X_{ik}d_{ikj}|d_{ikj} = 1]$.

Next, we aggregate up to the type-level:

$$\begin{aligned} \sum_l \left\{ Pr\left\{t_l^W >_{t_i^M} t_j^W\right\} \bar{X}_{t_i^M t_l^W t_j^W}\right\} + \sum_l \left\{ Pr\left\{t_l^M >_{t_j^W} t_i^M\right\} \bar{X}_{t_l^M t_j^W t_i^M}\right\} \\ (10) \qquad \qquad \qquad \geq |t_j^W| |t_i^M| (1 - E[X_{ij}]). \end{aligned}$$

Here $\bar{X}_{t_i^M t_l^W t_j^W} \equiv \sum_{k \in t_l^W} \sum_{i \in t_i^M} \sum_{j \in t_j^W} \bar{X}_{ikj}$ and $\bar{X}_{t_l^M t_j^W t_i^M} \equiv \sum_{j \in t_i^M} \sum_{j \in t_j^W} \sum_{i \in t_i^M} \bar{X}_{kji}$. In the above inequality, only the $|t_j^W|$ and $|t_i^M|$ are observed, but nothing else. This is of little use empirically.

On the other hand, because $d_{ijk} \geq 0$, for all (i, j, k) , we also have

$$\begin{aligned} E(X_{ik}d_{ikj}) &= E(X_{ik}d_{ikj}|d_{ikj} = 1)Pr(d_{ikj} = 1) \leq E(X_{ik}) \\ &\Rightarrow \sum_{k \in t_l^W} E(X_{ik}d_{ikj}|d_{ikj} = 1)Pr(d_{ikj} = 1) \leq \sum_{k \in t_l^W} E(X_{ik}) \\ &\Leftrightarrow Pr(t_l^W >_i j) \sum_{k \in t_l^W} \bar{X}_{ikj} \leq \sum_{k \in t_l^W} E(X_{ik}) \\ (11) \quad &\Rightarrow \sum_{i \in t_i^M} Pr(t_l^W >_i j) \sum_{k \in t_l^W} \bar{X}_{ikj} \leq \sum_{i \in t_i^M} \sum_{k \in t_l^W} E(X_{ik}) \\ &\Leftrightarrow Pr(t_l^W >_{t_i^M} j) \sum_{i \in t_i^M} \sum_{k \in t_l^W} \bar{X}_{ikj} \leq X_{t_i^M t_l^W} \\ &\Rightarrow Pr(t_l^W >_{t_i^M} t_j^W) \sum_{j \in t_j^W} \sum_{i \in t_i^M} \sum_{k \in t_l^W} \bar{X}_{ikj} \leq |t_j^W| X_{t_i^M t_l^W} \\ &\Leftrightarrow Pr(t_l^W >_{t_i^M} t_j^W) \bar{X}_{t_i^M t_l^W t_j^W} \leq |t_j^W| X_{t_i^M t_l^W} \end{aligned}$$

Combining inequalities (10) and (11), we get

$$\sum_l |t_j^W| X_{t_i^M t_l^W} + \sum_l |t_i^M| X_{t_l^M t_j^W} \geq |t_j^W| |t_i^M| (1 - E[X_{ij}])$$

By the equalities $\sum_l X_{t_i^M t_l^W} = |t_i^M|$ and $\sum_l X_{t_l^M t_j^W} = |t_j^W|$, the above reduces to

$$2 |t_i^M| |t_j^W| \geq |t_i^M| |t_j^W| (1 - E[X_{ij}]) \Rightarrow 2 \geq (1 - E[X_{ij}])$$

which is trivially satisfied.

REFERENCES

- Andrews, D., Berry, S., and Jia, P. (2004). Confidence regions for parameters in discrete games with multiple equilibria, with an application to discount chain store location. mimeo., Yale University.
- Beresteanu, A., Molchanov, I., and Molinari, F. (2011). Sharp identification regions in models with convex moment predictions. *Econometrica*, 79(6):1785–1821.
- Beresteanu, A. and Molinari, F. (2008). Asymptotic properties for a class of partially identified models. *Econometrica*, 76:763–814.
- Chernozhukov, V., Hong, H., and Tamer, E. (2007). Estimation and confidence regions for parameter sets in econometric models. *Econometrica*, 75:1234–1275.
- Choo, E. and Siow, A. (2006). Who marries whom and why. *Journal of Political Economy*, 114(1):175–201.
- Ciliberto, F. and Tamer, E. (2009). Market Structure and Multiple Equilibria in Airline Markets. *Econometrica*, 77:1791–1828.
- Dagsvik, J. K. (2000). Aggregation in matching markets. *International Economic Review*, 41(1):27–57.
- Echenique, F., Lee, S., Shum, M., and Yenmez, M. (2013). The revealed preference theory of stable and extremal stable matchings. *Econometrica*, 81:153–171.
- Fox, J. (2010). Identification in matching games. *Quantitative Economics*, 1(2):203–254.
- Galichon, A. and Salanie, B. (2009). Matching with trade-offs: Revealed preferences over competing characteristics. mimeo., Sciences Po.
- Pakes, A., Porter, J., Ho, K., and Ishii, J. (2007). Moment inequalities and their application. Manuscript, Harvard University.
- Romano, J. and Shaikh, A. (2010). Inference for the identified set in partially identified econometric models. *Econometrica*, 78:169–211.
- Shapley, L. and Shubik, M. (1971). The assignment game I: The core. *International Journal of Game Theory*, 1(1):111–130.
- Uetake, K. and Watanabe, Y. (2012). A note on estimation of two-sided matching models. *Economics Letters*.